

TITLE OF THE INVENTION

GENES ENCODING UMP KINASE, METHODS FOR PURIFYING UMP KINASE AND
METHODS OF CHARACTERIZING UMP KINASE

5

BACKGROUND OF THE INVENTION

The phosphorylation of UMP and CMP in eucaryotes is carried out by a single polypeptide. UMP/CMP kinases from *Saccharomyces cerevisiae*, *Dictyostelium discoideum*, *Arabidopsis thaliana* or pig muscle resemble adenylate kinase from muscle cytosol [Müller-Dieckmann, 1994 #717; Scheffzek, 1996 #750; Zhou, 1998 #1538; Okajima, 1995 #276; Dreusicke, 1988 #22]. Enteric bacteria contain separate UMP and CMP kinases and mutants defective in the gene encoding UMP kinase (*pyrH/smbA*) and CMP kinase (*mssA/cmk*) from *E. coli* or *Salmonella typhimurium* were isolated and characterized many years ago [Ingraham, 1972 #681; Beck, 1974 #636; Piérard, 1976 #735]. The recombinant UMP- and CMP kinases from *E. coli* were characterized in much detail [Serina, 1995 #951; Serina, 1996 #1058; Bucurenci, 1998 #985; Landais, 1999 #1080; Bucurenci, 1996 #984; Briozzo, 1998 #1651]. Thus, the bacterial UMP kinase is a homohexamer whose primary structure is divergent from that of other nucleoside monophosphate (NMP) kinases. The enzyme has an absolute specificity for UMP as substrate and is controlled allosterically by GTP (activator) and UTP (inhibitor) [Serina, 1995 #951]. CMP kinase from *E. coli* is a monomer which acts preferentially on CMP and dCMP [Bucurenci, 1996 #984]. Although the enzyme has little overall sequence identity with other known NMP kinases, it has in common with these enzymes a central parallel β -sheet, the strand of which are connected by α -helices. A property which is unique to the bacterial CMP kinase is a 40-residue insert situated within the CMP binding site and consisting of a three-stranded antiparallel β -sheet and two α -helices [Briozzo, 1998 #1651].

Attempts in the past to isolate a specific UMP kinase from *B. subtilis* failed. It was suggested that phosphorylation of UMP in this bacterium is accomplished by a CMP kinase with a broader specificity for pyrimidine nucleotides than the enzyme from *E. coli* [Waleh,

1976 #278]. The deleterious effect of disruption of *cmk/jofC* gene in *B. subtilis* [Sorokin, 1995 #277], was in line with this interpretation. Thanks to the genome sequencing programs, the *pyrH* gene was identified in all investigated bacteria, including *B. subtilis*. On the other hand, the *pyrH* gene from *Lactococcus lactis*, a bacterium similar to *B. subtilis* in the metabolism of pyrimidine nucleotides, complements a temperature sensitive *pyrH* mutation in *E. coli* demonstrating the ability of the encoded protein to synthesize UDP [Wadskov-Hansen, 2000 #1878].

These observations reopened the question of the role played by UMP kinase in the metabolism of *B. subtilis* and in gram positive organisms in general, and prompted us to clone the *pyrH* gene from *B. subtilis* and to examine the structural and catalytic properties of the recombinant protein. A striking characteristic of *B. subtilis* UMP kinase in comparison with the *E. coli* enzyme is its very low activity in the absence of GTP. On the other hand, the enzyme is unstable in the absence of UTP both in crude extract or under purified form. Antibodies against the recombinant UMP kinase identified the enzyme in the *B. subtilis* proteome, and immunoelectron microscopy confirmed the peripheral distribution of UMP kinase in this organism which extends our previous observations on *E. coli* enzyme.

In addition, we have successfully isolated and characterized clones of the *pyrH/subA* gene which encodes UMP kinase from *M. tuberculosis* and *H. influenzae*.

SUMMARY OF THE INVENTION

Accordingly, an object of the present invention is to identify and characterize specific genes which encode UMP kinase. Characterisation includes determination of the sensitivity to nucleotides and nucleotide analogs. For example, activation in the presence of UTP and inhibition in the presence of GTP.

The pathway leading to synthesis of UMP in prokaryotes is also present in *Bacillus subtilis*, a gram positive bacterium whose whole genome sequence was reported three years ago. A question still waiting an answer was related to the existence of an active UMP kinase in this organism, as attempts to isolate the enzyme were unsuccessful. The gene encoding the UMP kinase (*pyrH/smbA*) is present in *B. subtilis*, and its open reading frame is transcribed *in vivo* into a functional enzyme of X aminoacid residues, (~0.15 % of total proteins of *B. subtilis*). UMP kinase from *B. subtilis* is extremely unstable in the absence of cofactors. UTP in millimolar concentrations increased significantly the stability of the protein in long-term storage either in bacterial extract or under purified form. The specific activity of the purified enzyme in the presence of GTP which acts as an activator is of 25 $\mu\text{mol}/\text{min}^{-1}/\text{mg}$ of protein⁻¹. Taking into consideration, the specific activity of UMP kinase from *E. coli* under identical conditions 150 $\mu\text{mol}/\text{min}^{-1}/\text{mg}$ of protein⁻¹ and the relative abundance of the enzyme 0.05 % of total *E. coli* proteins it might be estimated that the rate of UMP phosphorylation in *B. subtilis* is approximately 50 % of that in *E. coli*. In the absence of GTP the activity of *B. subtilis* UMP kinase is twenty times lower, indicating the major role of this nucleotide in controlling catalysis both *in vitro* and *in vivo*. Activation by GTP is specific: contrary to *E. coli* UMP kinase which is also activated by GMP, the *B. subtilis* enzyme is insensitive to the latter nucleotide. Only dGTP and GMP-PNP can activate at significant rates the enzyme from *B. subtilis*. UTP inhibits the UMP kinase from *B. subtilis*, with a lower affinity than that shown towards the *E. coli* UMP kinase. Antibodies directed against the recombinant enzyme demonstrated the peripheral distribution of UMP kinase in *B. subtilis* extending our previous observations on the enzyme from *Escherichia coli*.

DETAILED DESCRIPTION OF THE INVENTION

The term "nucleic acid molecule" as used herein means RNA or DNA, including cDNA, single or double stranded, and linear or covalently closed molecules. A nucleic acid molecule may also be genomic DNA corresponding to the entire gene or a substantial portion thereof or to fragments and derivatives thereof. The nucleotide sequence may correspond to the naturally occurring nucleotide sequence or may contain single or multiple nucleotide substitutions, deletions and/or additions including fragments thereof. All such variations in the nucleic acid molecule retain the ability to encode a biologically active protein when expressed in the appropriate host or an enzymatically active fragment thereof. The nucleic acid molecule of the present invention may comprise solely the nucleotide sequence encoding a protein or may be part of a larger nucleic acid molecule that extends to the gene for the protein. The non-protein encoding sequences in a larger nucleic acid molecule may include vector, promoter, terminator, enhancer, replication, signal sequences, or non-coding regions of the gene.

Those nucleotide sequences which are substantially identical to those specifically disclosed are included in the present invention. Such sequences are those which hybridize to each other under stringent conditions and encode UMP kinase. Stringent conditions are sequence dependent and will be different in different circumstances. Generally, stringent conditions are selected to be about 5° C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. Typically, stringent conditions will be those in which the salt concentration is about 0.02 molar at pH8 and a temperature of approximately 60°C. These methods and others known in the art are described in Sambrook et al, Molecular Cloning : A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY (1989).

Nucleotide sequences are also substantially identical for purposes of this application when the polypeptides which they encode are substantially identical. Thus, where one nucleic acid sequence encodes essentially the same polypeptide as a second nucleic acid sequence,

the two nucleic acid sequences are substantially identical, even if they would not hybridize under stringent conditions due to silent substitutions permitted by the genetic code (see, Darnell et al. (1990) Molecular Cell Biology, Second Edition Scientific American Books W. H. Freeman and Company New York for codon degeneracy and the genetic code).

5 The conditions for culturing the microorganisms can be chosen from those which are preferable for their growth. Any natural- and synthetic-nutrient culture media can be used for culturing the microorganisms used in the present process as long as the microorganisms can grow therein and produce the present enzyme. The carbon sources used in the present invention are those which can be utilized by the microorganisms; for example, saccharides
10 such as maltose, trehalose, dextrins, and starches, and natural substances which contain saccharides such as molasses and yeast extracts can be used. The concentration of these carbon sources contained in the culture media is chosen depending on their types. The nitrogen sources used in the present invention are, for example, inorganic nitrogen-containing compounds such as ammonium salts and nitrates, and organic nitrogen-containing compounds
15 such as urea, corn steep liquor, casein, peptone, yeast extract, and meat extract. If necessary, inorganic compounds, for example, salts of calcium, magnesium, potassium, sodium, phosphoric acid, manganese, zinc, iron, copper, molybdenum, and cobalt can be used in the present invention.

After culturing the microorganisms, the present enzyme can be collected from the
20 cultures. Because the enzyme activity may be generally present intracellularly, intact and processed cells can be obtained as crude enzymes. Whole cultures can be also used as crude enzymes. Conventional solid-liquid separation methods can be used to separate cells and nutrient culture media; for example, methods to directly centrifuge the cultures, those to filtrate the cultures after adding filter aids to the cultures or after pre-coating, and those to
25 filter the cultures using membranes such as plain filters and hollow fibers can be used. The intact and processed cells per se can be used as crude enzymes, and if necessary, they can be prepared into partially purified enzymes.

The types of the processed cells include protein fractions of cells, immobilized substances of the intact and processed cells, and cells which were dried, lyophilized, and

treated with surfactants, enzymes, ultrasonication, mechanical grinding, and mechanical pressure.

Expression vectors for use in prokaryotic host cells generally comprise one or more phenotypic selectable marker genes. A phenotypic selectable marker gene is, for example, a gene encoding a protein that confers antibiotic resistance or that supplies an autotrophic requirement. Examples of useful expression vectors for prokaryotic host cells include those derived from commercially available plasmids such as the cloning vector pBR322. pBR322 contains genes for ampicillin and tetracycline resistance and thus provides simple means for identifying transformed cells. To construct an expression vector using pBR322, an appropriate promoter and a DNA sequence are inserted into the pBR322 vector.

Promoter sequences commonly used for recombinant prokaryotic host cell expression vectors include β -lactamase (penicillinase), lactose promoter system (Chang *et al.*, Nature 275:615, (1978); and Goeddel *et al.*, Nature 281:544, (1979)), tryptophan (trp) promoter system (Goeddel *et al.*, Nucl. Acids Res. 8:4057, (1980)), and tac promoter (Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Cold Spring Harbor Laboratory, p. 412 (1982)).

The proteins of the present invention may, when beneficial, be expressed as a fusion protein that has the protein attached to a fusion segment. The fusion segment often aids in protein purification, *e.g.*, by permitting the fusion protein to be isolated and purified by affinity chromatography. Fusion proteins can be produced by culturing a recombinant cell transformed with a fusion nucleic acid sequence that encodes a protein including the fusion segment attached to either the carboxyl and/or amino terminal end of the protein. Preferred fusion segments include, but are not limited to, glutathione-S-transferase, β -galactosidase, a poly-histidine segment capable of binding to a divalent metal ion, and maltose binding protein.

According to the present invention, isolated and purified UMP kinase may be produced by the recombinant expression systems described above. The method comprises

culturing a host cell transformed with an expression vector comprising a DNA sequence that encodes the protein under conditions sufficient to promote expression of the protein. The protein is then recovered from culture medium or cell extracts, depending upon the expression system employed. As is known to the skilled artisan, procedures for purifying a recombinant protein will vary according to such factors as the type of host cells employed and whether or not the recombinant protein is secreted into the culture medium. When expression systems that secrete the recombinant protein are employed, the culture medium first may be concentrated. Following the concentration step, the concentrate can be applied to a purification matrix such as a gel filtration medium. Alternatively, an anion exchange resin can be employed, *e.g.*, a matrix or substrate having pendant diethylaminoethyl (DEAE) groups. The matrices can be acrylamide, agarose, dextran, cellulose, or other types commonly employed in protein purification. Also, a cation exchange step can be employed. Suitable cation exchangers include various insoluble matrices comprising sulfoethyl or carboxymethyl groups. Further, one or more reversed-phase high performance liquid chromatography (RP-HPLC) steps employing hydrophobic RP-HPLC media (*e.g.*, silica gel having pendant methyl or other aliphatic groups) can be employed to further purify the protein. Some or all of the foregoing purification steps, in various combinations, are well known in the art and can be employed to provide an isolated and purified recombinant protein. These and other methods are disclosed in Sambrook et al, Molecular Cloning : A Laboratory Manual, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, 1989.

In the methods of detecting, characterizing or stabilizing a UMP kinase, GTP and UTP are typically used. However, other nucleotides and/or nucleotide analogs may be used in accordance with the disclosure in the present application. Analogs of GTP, UTP, TTP, ATP, and CTP are known in the art.

EXAMPLES

Chemicals. Nucleotides, restriction enzymes, T4 DNA ligase T7 DNA polymerase and coupling enzymes were from Roche-Diagnostics or from Sigma. NDP kinase from
5 *Dictyostelium discoideum* (2000 U/mg of protein) was kindly provided by M. Véron.

Bacterial Strains, Plasmids, Growth Conditions and DNA Manipulations. General DNA manipulations were performed as described by Sambrook et al. [Sambrook, 1989 #66]. The
10 *pyrH* gene from *B. subtilis* was amplified by polymerase chain reaction using chromosomal DNA from the strain 168 () as the matrix. The product was inserted between the *NdeI* and *XhoI* restriction sites of plasmid pET24a (Novagen). The resulting plasmid (pSL13) was introduced into strain BL21(DE3)/pDIA17 () to overproduce the UMP kinase. Recombinant strain was grown in 2YT medium supplemented with antibiotics to an optical density of 1 at 600 nm. Then overproduction was triggered by isopropyl- β -D-thiogalactoside induction (1
15 mM final concentration) for 3 h, then bacteria were harvested by centrifugation.

Purification of UMP kinase and activity assays. *E. coli* overproducing the UMP kinase from *B. subtilis* was disrupted by sonication in 50 mM Tris-HCl (pH 7.4) and 2 mM UTP. The bacterial extract was heated for 10 min at 65°C, then the precipitated proteins were
20 removed by centrifugation at 10,000 g for 30 min. The supernatant was concentrated by ultrafiltration to about 10 mg of protein/ml, then applied to a Sephacryl S-300 HR column (2.5 x 110 cm) equilibrated with 50 mM Tris-HCl (pH 7.4), 0.1 M NaCl and 2 mM UTP at a flow rate of 10 ml/h. The peak fraction containing over 95 % pure UMP kinase was concentrated again to 10 mg of protein/ml. Fractions of 1 ml of protein solution were sampled
25 in Eppendorf tubes and stored at different temperatures between 20°C and -80°C. The UMP kinase activity was determined at 30°C using a coupled spectrophotometric assay (0.5 ml final volume) on an Eppendorf PCP6121 photometer [Blondin, 1994 #977]. The "standard" reaction medium contained 50 mM Tris-HCl (pH 7.4), 50 mM KCl, 2 mM $MgCl_2$, 1 mM phosphoenolpyruvate, 0.2 mM NADH, 2 mM ATP, 0.5 mM GTP and 2 units each of lactate
30 dehydrogenase, pyruvate kinase and NDP kinase. The crude or pure preparation of UMP

kinase was then added, followed two minutes later by 1.3 mM UMP. The decrease in absorbance at 334 nm (between 0.03 and 0.3/min) was then recorded and corrected for secondary reactions, occurring in the absence of UMP. One unit of UMP kinase corresponds to 1 μ mol of product formed per min.

5

Immunochemical methods. Anti-UMP kinase sera were obtained by immunizing rabbits with 250 μ g of purified recombinant protein at 12 days intervals. After four injections, the rabbits were bled and polyclonal response tested by ELISA. Immune sera were adsorbed against an *E. coli* sonicate to improve the signal-to-noise ratio. Western blotting was performed after SDS-polyacrylamide gel electrophoresis (SDS-PAGE) by transfer of proteins on a nitrocellulose membrane, followed by treatment with 1:1000 dilution of polyclonal sera and alkaline phosphatase-conjugated anti-rabbit immunoglobulins. Alkaline phosphatase activity was revealed by using the nitroblue tetrazolium-5-bromo-4-chloro-3-indolylphosphate dye system.

10

11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
841
842
843
844
845
846
847
848
849
850
851
852
853
854
855
856
857
858
859
860
861
862
863
864
865
866
867
868
869
870
871
872
873
874
875
876
877
878
879
880
881
882
883
884
885
886
887
888
889
890
891
892
893
894
895
896
897
898
899
900
901
902
903
904
905
906
907
908
909
910
911
912
913
914
915
916
917
918
919
920
921
922
923
924
925
926
927
928
929
930
931
932
933
934
935
936
937
938
939
940
941
942
943
944
945
946
947
948
949
950
951
952
953
954
955
956
957
958
959
960
961
962
963
964
965
966
967
968
969
970
971
972
973
974
975
976
977
978
979
980
981
982
983
984
985
986
987
988
989
990
991
992
993
994
995
996
997
998
999
1000
1001
1002
1003
1004
1005
1006
1007
1008
1009
1010
1011
1012
1013
1014
1015
1016
1017
1018
1019
1020
1021
1022
1023
1024
1025
1026
1027
1028
1029
1030
1031
1032
1033
1034
1035
1036
1037
1038
1039
1040
1041
1042
1043
1044
1045
1046
1047
1048
1049
1050
1051
1052
1053
1054
1055
1056
1057
1058
1059
1060
1061
1062
1063
1064
1065
1066
1067
1068
1069
1070
1071
1072
1073
1074
1075
1076
1077
1078
1079
1080
1081
1082
1083
1084
1085
1086
1087
1088
1089
1090
1091
1092
1093
1094
1095
1096
1097
1098
1099
1100
1101
1102
1103
1104
1105
1106
1107
1108
1109
1110
1111
1112
1113
1114
1115
1116
1117
1118
1119
1120
1121
1122
1123
1124
1125
1126
1127
1128
1129
1130
1131
1132
1133
1134
1135
1136
1137
1138
1139
1140
1141
1142
1143
1144
1145
1146
1147
1148
1149
1150
1151
1152
1153
1154
1155
1156
1157
1158
1159
1160
1161
1162
1163
1164
1165
1166
1167
1168
1169
1170
1171
1172
1173
1174
1175
1176
1177
1178
1179
1180
1181
1182
1183
1184
1185
1186
1187
1188
1189
1190
1191
1192
1193
1194
1195
1196
1197
1198
1199
1200
1201
1202
1203
1204
1205
1206
1207
1208
1209
1210
1211
1212
1213
1214
1215
1216
1217
1218
1219
1220
1221
1222
1223
1224
1225
1226
1227
1228
1229
1230
1231
1232
1233
1234
1235
1236
1237
1238
1239
1240
1241
1242
1243
1244
1245
1246
1247
1248
1249
1250
1251
1252
1253
1254
1255
1256
1257
1258
1259
1260
1261
1262
1263
1264
1265
1266
1267
1268
1269
1270
1271
1272
1273
1274
1275
1276
1277
1278
1279
1280
1281
1282
1283
1284
1285
1286
1287
1288
1289
1290
1291
1292
1293
1294
1295
1296
1297
1298
1299
1300
1301
1302
1303
1304
1305
1306
1307
1308
1309
1310
1311
1312
1313
1314
1315
1316
1317
1318
1319
1320
1321
1322
1323
1324
1325
1326
1327
1328
1329
1330
1331
1332
1333
1334
1335
1336
1337
1338
1339
1340
1341
1342
1343
1344
1345
1346
1347
1348
1349
1350
1351
1352
1353
1354
1355
1356
1357
1358
1359
1360
1361
1362
1363
1364
1365
1366
1367
1368
1369
1370
1371
1372
1373
1374
1375
1376
1377
1378
1379
1380
1381
1382
1383
1384
1385
1386
1387
1388
1389
1390
1391
1392
1393
1394
1395
1396
1397
1398
1399
1400
1401
1402
1403
1404
1405
1406
1407
1408
1409
1410
1411
1412
1413
1414
1415
1416
1417
1418
1419
1420
1421
1422
1423
1424
1425
1426
1427
1428
1429
1430
1431
1432
1433
1434
1435
1436
1437
1438
1439
1440
1441
1442
1443
1444
1445
1446
1447
1448
1449
1450
1451
1452
1453
1454
1455
1456
1457
1458
1459
1460
1461
1462
1463
1464
1465
1466
1467
1468
1469
1470
1471
1472
1473
1474
1475
1476
1477
1478
1479
1480
1481
1482
1483
1484
1485
1486
1487
1488
1489
1490
1491
1492
1493
1494
1495
1496
1497
1498
1499
1500
1501
1502
1503
1504
1505
1506
1507
1508
1509
1510
1511
1512
1513
1514
1515
1516
1517
1518
1519
1520
1521
1522
1523
1524
1525
1526
1527
1528
1529
1530
1531
1532
1533
1534
1535
1536
1537
1538
1539
1540
1541
1542
1543
1544
1545
1546
1547
1548
1549
1550
1551
1552
1553
1554
1555
1556
1557
1558
1559
1560
1561
1562
1563
1564
1565
1566
1567
1568
1569
1570
1571
1572
1573
1574
1575
1576
1577
1578
1579
1580
1581
1582
1583
1584
1585
1586
1587
1588
1589
1590
1591
1592
1593
1594
1595
1596
1597
1598
1599
1600
1601
1602
1603
1604
1605
1606
1607
1608
1609
1610
1611
1612
1613
1614
1615
1616
1617
1618
1619
1620
1621
1622
1623
1624
1625
1626
1627
1628
1629
1630
1631
1632
1633
1634
1635
1636
1637
1638
1639
1640
1641
1642
1643
1644
1645
1646
1647
1648
1649
1650
1651
1652
1653
1654
1655
1656
1657
1658
1659
1660
1661
1662
1663
1664
1665
1666
1667
1668
1669
1670
1671
1672
1673
1674
1675
1676
1677
1678
1679
1680
1681
1682
1683
1684
1685
1686
1687
1688
1689
1690
1691
1692
1693
1694
1695
1696
1697
1698
1699
1700
1701
1702
1703
1704
1705
1706
1707
1708
1709
1710
1711
1712
1713
1714
1715
1716
1717
1718
1719
1720
1721
1722
1723
1724
1725
1726
1727
1728
1729
1730
1731
1732
1733
1734
1735
1736
1737
1738
1739
1740
1741
1742
1743
1744
1745
1746
1747
1748
1749
1750
1751
1752
1753
1754
1755
1756
1757
1758
1759
1760
1761
1762
1763
1764
1765
1766
1767
1768
1769
1770
1771
1772
1773
1774
1775
1776
1777
1778
1779
1780
1781
1782
1783
1784
1785
1786
1787
1788
1789
1790
1791
1792
1793
1794
1795
1796
1797
1798
1799
1800
1801
1802
1803
1804
1805
1806
1807
1808
1809
1810
1811
1812
1813
1814
1815
1816
1817
1818
1819
1820
1821
1822
1823
1824
1825
1826
1827
1828
1829
1830
1831
1832
1833
1834
1835
1836
1837
1838
1839
1840
1841
1842
1843
1844
1845
1846
1847
1848
1849
1850
1851
1852
1853
1854
1855
1856
1857
1858
1859
1860
1861
1862
1863
1864
1865
1866
1867
1868
1869
1870
1871
1872
1873
1874
1875
1876
1877
1878
1879
1880
1881
1882
1883
1884
1885
1886
1887
1888
1889
1890
1891
1892
1893
1894
1895
1896
1897
1898
1899
1900
1901
1902
1903
1904
1905
1906
1907
1908
1909
1910
1911
1912
1913
1914
1915
1916
1917
1918
1919
1920
1921
1922
1923
1924
1925
1926
1927
1928
1929
1930
1931
1932
1933
1934
1935
1936
1937
1938
1939
1940
1941
1942
1943
1944
1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025
2026
2027
2028
2029
2030
2031
2032
2033
2034
2035
2036
2037
2038
2039
2040
2041
2042
2043
2044
2045
2046
2047
2048
2049
2050
2051
2052
2053
2054
2055
2056
2057
2058
2059
2060
2061
2062
2063
2064
2065
2066
2067
2068
2069
2070
2071
2072
2073
2074
2075
2076
2077
2078
2079
2080
2081
2082
2083
2084
2085
2086
2087
2088
2089
2090
2091
2092
2093
2094
2095
2096
2097
2098
2099
2100
2101
2102
2103
2104
2105
2106
2107
2108
2109
2110
2111
2112
2113
2114
2115
2116
2117
2118
2119
2120
2121
2122
2123
2124
2125
2126
2127
2128
2129
2130
2131
2132
2133
2134
2135
2136
2137
2138
2139
2140
2141
2142
2143
2144
2145
2146
2147
2148
2149
2150
2151
2152
2153
2154
2155
2156
2157
2158
2159
2160
2161
2162
2163
2164
2165
2166
2167
2168
2169
2170
2171
2172
2173
2174
2175
2176
2177
2178
2179
2180

Immunoelectron microscopy. Bacteria were fixed with 4% formaldehyde (freshly made from paraformaldehyde) and 0.2% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.4), for one hour at 4°C. The cell pellets were rinsed with cacodylate buffer, then treated with 0.5% aqueous uranyl acetate solution [Benichou, 1990 #819], followed by a final rinse in distilled water. Bacteria were embedded in 2% agarose (type IX, Sigma). Small blocks were embedded in Unicryl by the PLT method and modified procedure, as described by Gounon and Rolland (15) [Gounon, 1999 #1880]. Ultrathin sections were collected on Formvar-carbon coated nickel grids. Sections were then incubated in the following solutions: PBS containing 50 mM NH_4Cl : 10 min; PBS containing 1% BSA and 1% normal goat serum [Brorson, 1997 #817]: 10 min; rabbit polyclonal anti-UMP kinase antisera (1/100 dilution), or mouse monoclonal anti-CMP kinase antibodies (100 $\mu\text{g/ml}$): one hour. Two washes (5 min each) were performed in PBS containing 0.1% BSA, then one wash in PBS. Incubations were for 45 min in a solution containing anti-rabbit or anti-mouse gold-conjugated immunoglobulin (5 nm or 10 nm particles, British Biocell Laboratories, Cardiff, UK), diluted 1/20 in PBS containing 0.01% fish skin gelatin (Sigma). Sections were washed once in PBS and three times in distilled water, then fixed for 2 min with 1% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.4), and finally rinsed with distilled water and dried. Optional counterstaining was performed by treating the sections with 2% aqueous uranyl acetate solution for 40 min, followed by a 3 min incubation in Millonig's lead tartrate solution [Millonig, 1961 #826]. Specimens were examined with a Philips CM12 electron microscope operating under standard conditions [Landais, 1999 #1080].

Other analytical procedures. Protein concentration was measured according to Bradford [Bradford, 1976 #68]. Ion spray mass spectra were recorded on a quadrupole mass spectrometer API-365 (Perkin-Elmer) equipped with an ion spray (nebulizer-assistant electrospray) source. The sample ($\sim 2 \text{ pmol} \cdot \mu\text{l}^{-1}$) dissolved in 20 % acetonitrile in water and 0.1 % HCOOH was delivered to the source at a flow rate of $5 \mu\text{l} \cdot \text{min}^{-1}$. SDS-PAGE was performed as described by Laemmli [Laemmli, 1970 #69]. The proteins bands from SDS-

PAGE were electroblotted into a Problott membrane filter (Applied Biosystems) and detected by staining in the Coomassie Blue. The N-terminal amino acid sequence of the protein from the excised band was determined by a protein sequencer (Applied Biosystems, Inc.). Fluorescence experiments were performed on a Perkin-Elmer LS-5B luminescence spectrometer thermostated at 25°C. Emission spectra of UMP kinase ($\lambda_{exc} = 295$ nm; band width = 5 nm) were recorded from 305 to 400 nm.

Results

Cloning, sequencing of *pyrH* gene from *B. subtilis* and complementation tests in *E. coli*.

The *pyrH* gene from *B. subtilis* was cloned by PCR into the expression vector pET24a, and
sequenced. The resulting ORF showed two differences when compared with the published
databank : one additional T at bp170, and one missing A at bp185. As a consequence, the
ORF of the *pyrH* gene displays a double frame-shift of 14 bp long stretch, resulting in four
amino acid residues change : ⁵⁷LeuTrpArgGly⁶⁰ instead of TyrGlyAlaGlu in the original
sequence. Harbored on high-copy number vectors, the *B. subtilis pyrH* gene complemented
the thermosensitive phenotype of strain KUR1244 (*pyrH88ts*) of *E. coli* indicating that it was
functional. Complementation experiments performed on strain MC4100-42-14 :40
(*car* : :*lacZpyrH42*), in which expression of the *car* : :*lacZ* fusion is repressed in the presence
of wild-type UMP kinase activity, showed that in high copy-number, the *pyrH* gene from
B. subtilis resulted in a significant repression of β -galactosidase activity.

Stability of recombinant UMP kinase in crude bacterial extracts.

UMP kinase from *B. subtilis* overproduced in strain BL21 (DE3)/pDIA17 (over 30% of total
E. coli proteins) was recovered in the supernatant after cell breakage in 50 mM Tris-HCl (pH
7.4) and centrifugation. The activity of recombinant UMP kinase declined rapidly upon
storage at either room temperature, +4°C or under frozen state. After 24 h, only 15% of initial
UMP kinase activity was recovered. EDTA, thiols, bovine serum albumin, or antiprotease
compounds alone or in mixture were ineffective. UTP (≥ 2 mM) stabilized considerably the
bacterial UMP kinase when stored at various temperatures. A mixture of ATP and GTP (2
mM each) was equally effective. UTP increased the thermal stability of bacterial UMP kinase

(Fig. 1), the half maximal inactivation being shifted from 42°C in the absence of UTP to over 70°C in the presence of nucleotide.

Purification and molecular characterization of recombinant UMP kinase from

5 *B. subtilis*.

UMP kinase from *B. subtilis* overproduced in strain BL21(DE3)/pDIA17 was purified as described under experimental procedures, i.e. a heating step followed by gel permeation chromatography (Fig. 2). The molecular mass of *B. subtilis* UMP kinase ($26,084.2 \pm 1.7$ Da), measured by ESI-MS was in agreement with that calculated (26,083 Da) from the sequence.

10 Gel permeation chromatography yielded a single symmetrical peak of protein consistent with an oligomeric enzyme (6 subunits/oligomer). Ultracentrifugation analysis by sedimentation equilibrium indicated that the dominant species (156 KDa) corresponded to the hexameric enzyme, eventought oligomers of higher molecular mass were also identified.

In contrast to *E. coli* UMP kinase, the *B. subtilis* enzyme was sensitive to trypsin digestion.

15 The inactivation of the bacterial enzyme at 30°C and in a 1/500 (w/w) trypsin/UMP kinase ratio followed a first order kinetics ($t_{1/2} =$ min). SDS-PAGE analysis of the digested enzyme showed accumulation of stable fragments. The truncated protein was still oligomeric as indicated by gel permeation chromatography. ESI-MS analysis of the partially cleaved protein indicated that the stable fragment ($18,905.8 \pm 0.5$ Da) is a C-terminal truncated form

20 (residues 1 to 174). A genetically engineered N-terminal His-tagged truncated form of enzyme was found essentially instable in the sonicated bacterial extract. The soluble fraction was purified by Ni-NTA chromatography. Gel permeation chromatography and appropriate molecular mass markers (lactate dehydrogenase, 140 KDa; creatine kinase, 82 KDa; and *E. coli* adenylate kinase, 27 KDa) indicated molecular mass of the recombinant protein of 120

25 KDa, consistent with a hexameric form. Urea induced denaturation of native and C-terminal truncated form of *B. subtilis* UMP kinase was monitored from the intrinsic fluorescence of

the single Trp residue (W58). As shown in Fig. the native intact UMP kinase irrespective of the presence or the absence of the His-tag exhibits upon excitation at 295 nm a fluorescence emission spectrum with maximum at nm which indicates that W58 is located in a hydrophobic environment, not exposed to the solvent. Urea over xM increased slightly the fluorescence maxima with shift to the red side of the emission spectrum. The mid point transition concentration of urea was xM. Under similar experimental conditions, the C-terminal truncated form of *B. subtilis* UMP kinase showed a significantly lower stability against denaturation by urea.

0 **Presence of the *pyrH* gene product in *B. subtilis* 168 wild-type strain.**

To determine the presence of *pyrH/smbA* gene product in the *B. subtilis* strain 168, the bacterial extract in 50 mM Tris-HCl (pH 7.4) and 2 mM UTP was heated at 65°C for 10 min, to inactivate ATPase activity interfering in the spectrophotometric assay. The specific activity of the protein in the crude extract under “ standard ” assay conditions (see Experimental Procedures) was of 0.016 U/mg protein. Since the purified recombinant protein has a specific activity of 26 U/mg of protein under identical experimental conditions, we might assume a protein abundance in *B. subtilis* extract between 0.06 and 0.1 %, a figure close to that found in *E. coli* where UMP kinase represents 0.05 % of total proteins. The UMP kinase was identified in the 2D-PAGE map of wild-type strain 168 by comigration with recombinant protein. Further unambiguous identification of enzyme was performed by 2D western blotting using rabbit polyclonal antibodies raised against the recombinant protein. The *B. subtilis* *pyrH* gene product migrates at $pI=$ and apparent molecular mass of KDa. Densitometric scanning of silver stained 2D gels indicated a protein abundance of 0.15 % which fitted reasonably well with that calculated from specific activities.

Kinetic properties of UMP kinase from *B. subtilis*.

Determination of UMP kinase activity with various nucleoside triphosphates and UMP at fixed concentrations (1 mM) indicated low specific activities, the maximal rate being with ATP. When NTPs were used in mixture, the highest specific activity was obtained with ATP and GTP, indicating an almost absolute requirement for GTP for expression of full catalytic activity dATP was as good as ATP as phosphate donor. When ATP concentration was varied in the absence or in the presence of GTP (0.5 mM) at a single concentration of UMP (1 mM), the apparent K_m for ATP was unusually high (0.9 mM). When ATP was constant (1 mM) the kinetics with variable concentrations of UMP was strongly dependant on GTP. Thus in absence of GTP, the rates were maximal at 50-70 μ M UMP (K_m for UMP \approx 8 μ M) to decline upon further increase in UMP. In the presence of GTP, the saturation was attained at 0.2 mM UMP. The apparent K_m for UTP was 30 μ M without inhibition by excess of UMP. GTP showed by far the most important effect, i. e. a ten fold activation with 2 mM ATP and 1 mM UTP, the half maximum activation being reach at 0.1 mM nucleoside triphosphate. dGTP was also effective but with lower affinities and extent of activation whereas GMP was totally ineffective. These results are in contrast with those obtained with *E. coli* UMP kinase, where GMP, cGMP and even guanosine exerted a significant activation. UTP antagonized the effect of GTP. In the absence, UTP decreased the reaction rate with an I_{50} value of approximately 60 μ M. At lower concentration of UMP (50 μ M) the I_{50} of UTP inhibition was decreased to 50 μ M.

Effect of uridine nucleotide analogs on UMP kinase from *B. subtilis*.

Molecular cloning and overexpression of UMP kinase from the three organisms

The *pyrH/smbA* gene encoding UMP kinase from *M. tuberculosis*, *B. subtilis* and *H. influenzae* was cloned by PCR amplification, using as template DNA isolated from the three strains employed for the whole genome sequencing.

a) The *pyrH* gene of *M. tuberculosis* was cloned using the PRO bacterial expression system developed by Clontech. The vector used is a derivative of PROTet where the recombinant protein is fused to an N-terminal 6xHN affinity tag. This system is optimized for use with TALON resins, which are cobalt-based IMAC (immobilized metal ion affinity chromatography) resins. UMP kinase expressed in DH5 α PRO *E. coli* strain was induced when bacteria reached an OD of 0.7 by addition of anhydrotetracycline (final concentration of 100 μ g/liter). Bacteria were collected after incubation at 37°C for 16 hours. The plasmid containing the *pyrH* gene from *M. tuberculosis* has been deposited at the CNCM on August 08, 2000 under the accession number I-2542.

b) The *pyrH* gene from *B. subtilis* and *H. influenzae* was inserted into the plasmid pET22b (Novagen). The resulting plasmids are under the control of a hybrid promoter/operator region constituted from the T7 promoter followed by the lac operator. They were introduced into the strain BL21 (DE3) of *E. coli* (Novagen) that is producing the T7RNA polymerase. The strains were grown in LB medium supplemented with kanamycin (30 mg/liter) and chloramphenicol (30 mg/liter). Induction was made with 1 mM IPTG. Three hours after induction bacteria were collected by centrifugation and stored at -20°C until use. The plasmid containing the *pyrH* gene from *B. subtilis* has been deposited at the CNCM on November 17, 2000 under the accession number I-2579. The plasmid containing the *pyrH* gene from *H. influenzae* *smbA* insert has been deposited at the CNCM on October 26, 2000 under the accession number I-2574.

Purification and storage of UMP kinase from the three organisms

a) The recombinant UMP kinase from *M. tuberculosis* was purified according to the CLONTECH's TALON Purification Kit. The protein was eluted by increasing imidazole concentration from 5 to 150 mM. The tagged UMP kinase exhibits the same properties as the wild-type UMP kinase (hexamer, activation by GTP, inhibition by UTP Tm of 67°C in the presence of UTP and 52°C in the absence of UTP). It is soluble at > 10 mg/ml in 50 mM phosphate buffer (pH 7.4) with 100 mM NaCl and stable at 4°C for at least three months.

b) The recombinant UMP kinases from *B. subtilis* and *H. influenzae* were purified using essentially the protocol described below. The bacteria were disrupted by sonication in 50 mM

Tris-HCl pH 7.4 and 2 mM UTP. The bacterial extract was heated for 10 min at 65°C (*B. subtilis*) or 70°C (*H. influenzae*) then the precipitated proteins were removed by centrifugation at 10000g for 30 min. The supernatant was concentrated by ultrafiltration then applied to a Sephacryl S-300 HR column (2.5 x 110 cm) equilibrated with 50 mM Tris-HCl pH 7.4, 0.1 mM NaCl, and 2 mM UTP at a flow rate of 10 ml/h. The peak fractions containing over 95 % pure UMP kinase were pooled and concentrated again to 10-15 mg of protein/ml. Protein solution was kept at -20°C.

c) Important notice for storage

Whereas UMP kinase from *M. tuberculosis* is stable in the absence of UTP the enzyme from the other two organisms (particularly that from *B. subtilis*) are unstable in the absence of the nucleotide, irrespective of the storage temperature (+4°C, -20°C or -80°C).

Activity assay

The UMP kinase activity was determined at 30°C using a coupled spectrophotometric assay (0.5 ml final volume) on an Eppendorf PCP6121 photometer. The "standard" reaction medium contained 50 mM Tris-HCl (pH 7.4), 50 mM KCl, 2 mM MgCl₂, 1 mM phosphoenolpyruvate, 0.2 mM NADH, 1 or 2 mM ATP, 0.5 mM GTP and 2 units each of lactate dehydrogenase, pyruvate kinase and NDP kinase. The pure preparation of UMP kinase was then added, followed by 0.3 or 1 mM UMP. The decrease in absorbance at 334 nm (between 0.03 and 0.3/min) was then recorded and corrected for secondary reactions occurring in the absence of UMP. One unit of UMP kinase corresponds to 1 µmol of product formed per min. The specific activities of the recombinant enzymes are the following:

- 30 units/mg of protein for UMP kinase from *M. tuberculosis* (1 mM ATP, 0.3 mM UMP).
- 25 units/mg of protein for UMP kinase from *B. subtilis* (2 mM ATP, 1 mM UMP).
- 45 units/mg of protein for UMP kinase from *H. influenzae* (1 mM ATP, 1 mM UMP).

It should be noticed that the degree of activation by GTP is variable (a factor of 2 for UMP kinase from *M. tuberculosis*, of 3 for *H. influenzae*, and between 10 and 20 for *B. subtilis*). There are also differences concerning the best activator: in the case of *H. influenzae* the highest activation and affinity for UMP kinase is shown by cGMP, maximal activation being attained at 20 µM nucleotide.

Other UMP kinases

Continuing our searches for soluble and stable UMP kinases we identified two other enzymes not yet fully characterized: *Streptococcus pneumoniae* (gram positive) and *Neisseria meningitidis* (gram negative). The enzyme from the first organism resembles that from *B. subtilis* (high factor of activation by GTP), the enzyme from the second resembles that of *E. coli* and *H. influenzae*.

The ORFs of the *pyrH* gene from the three organisms

They are available in the gene databank. It should be mentioned, however, that we found two differences with the published sequence of *B. subtilis pyrH* gene: one additional T at bp 170, and one missing A at bp185. As a consequence, the ORF of the *pyrH* gene displays a double frame-shift of 14 bp long stretch, resulting in four amino acid residues change: ⁵⁷LeuTrpArgGly⁶⁰ instead of TyrGlyAlaGlu in the original deduced protein sequence. The *B. subtilis pyrH* gene is shown SEQ ID NO:1, the sequence of the *H. influenzae pyrH* gene is shown in SEQ ID NO:2, and the sequence of the *M. tuberculosis* gene is shown in SEQ ID NO:3.

All patents and publications mentioned herein are incorporated herein by reference to the extent allowed by law for the purpose of describing and disclosing the proteins, enzymes, vectors, host cells, and methodologies reported therein that might be used with the present invention.

References:

1. L. Serina, C. Blondin, E. Krin, O. Sismeiro, A. Danchin, H. Sakamoto, A.-M. Gilles & O. Bâzu. (1995). *Escherichia coli* UMP-kinase, a member of the aspartofinase family, is a hexamer regulated by guanine nucleotides and UTP. *Biochemistry*, 34, 5066-5074/
2. L. Serina, N. Bucurenci, A.-M. Gilles, W.K. Surewicz, H. Fabian, H.H. Mantsch, M. Takahashi, I. Petrescu, G. Batelier & O. Bâzu. (1996). Structural properties of UMP kinase from *Escherichia coli*: modulation of protein solubility by pH and UTP. *Biochemistry*, 35, 7003-7011.
3. N. Bucurenci, L. Serina, C. Zaharia, S. Landais, A. Danchin & O. Bâzu. (1998). Mutational analysis of UMP kinase from *Escherichia coli*. *J. Bacteriol.*, 180, 473-477.
4. S. Landais, P. Gounon, C. Laurent-Winter, J.-C. Mazié, A. Danchin, O. Bâzu & H. Sakamoto. (1999). Immunochemical analysis of UMP kinase from *Escherichia coli*. *J. Bacteriol.* 181, 833-840.
5. O. Bâzu, A.-M. Gilles & P. Briozzo. (1999). Bacterial UMP- and CMP kinases are distinct and structurally unrelated entities. *Paths to Pyrimidines* 7, 86-95.